



IFWO

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:47

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw

**ENTERED**

```

4 <110> APPLICANT: Kosan Biosciences, Inc.
5   Julien, Bryan
7 <120> TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE
8   INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS
9   BACTERIOPHAGE MX9
11 <130> FILE REFERENCE: 300622009940
13 <140> CURRENT APPLICATION NUMBER: US 10/645,818
14 <141> CURRENT FILING DATE: 2003-08-20
16 <150> PRIOR APPLICATION NUMBER: US 60/405,196
17 <151> PRIOR FILING DATE: 2002-08-21
19 <160> NUMBER OF SEQ ID NOS: 20
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1647
25 <212> TYPE: DNA
26 <213> ORGANISM: Bacteriophage MX9
28 <400> SEQUENCE: 1
29 gtggcgctca ggggtgcgtc ggacgccact accaaccctt ctcgacttgt gcagtcctgc 60
30 gccgccggcc cgcgtgcgac tccgtggggg gtcagtgcgt cgtggtagct gctagggcgt 120
31 acagcaacgg gggagtacat cgtgagtagc gacgcggcga agaagggccca tccaatggca 180
32 actgcggcgg agcgggttgc gacgtcacca atcgacgtca acgctctggc gctggagggtg 240
33 gcccggttgg tggccctcca gcagcaaagt gcgacgcgcg catcgtccgg ccgcactttc 300
34 ggcgcggtgg cggatgactg gctcatcact gaggccaagc gcctcgtgtg ccccgacaat 360
35 gacgcgccgc atcttcgcca tatggaggcg ctctggggca tgacggatgt ggagctcacg 420
36 ccgcgcgctg tgaaggcgca cctggcgagg cttctcaagc cagagggggc gctgagcgca 480
37 gccaccgtca ataagggtgc ctctaccggc aagcgcacat tcaaggcggc gcaaataaac 540
38 ggcgagtggg gcccggtgaa tcctttcgcc gtgctcgacc gcgaaaaaga ggcgagggcc 600
39 gacgcctcca cgtgacggc agcggagtgc cgggcgggtg tcccgcactt ccgcggggac 660
40 cggcgccgcg agtttctctt ccaggtcttt ctggggccac gcccgcggcg agagaaggcg 720
41 ctctcaagg aagatgtgga cgtcgaggcg cgcaccgtca ttttcggcg cagcaatgga 780
42 cgagacacga caaagacggg acgcgagcgt cgcgtgccgg tgccggatga gttgtggccc 840
43 gtgctcctcg atgcgatgca ggccagtcg tctgacctcg ttttcccgaa cgcgaagggt 900
44 gagaggcagc gcgcagacac gaagatgacg cgcgtgctgc gcaactgcgt atccgcggct 960
45 ggtgtcgtgg tgggctggga ttacatctgc cgcacgcagg gctgcggcta ccgagatgtg 1020
46 cagtctggtg gcgcgcgcca ggagcgtcgg tgccccgcct gcgacaagcg catgtggggc 1080
47 agtggtcgcc ccaaaccgcg cgtctggtac gggctccgtc acaccgcggc gacactgcac 1140
48 aggaaggcgg gctgcgaccc gctcgtcatc aagctcgtgc tggggcatgc ggctgtcgac 1200
49 accacggacg acgtgtacac gcacctcgac gaggactact gccgcgcga acttaacaag 1260
50 ttgtcgctga agggcccgcc gccaccact actaccagg gaggaagtga cggcgccct 1320
51 gactcaggag gcaacaccta cgggtgaagga ggcacatgc acggattggg agatttgag 1380
52 catcaccggg cgagagcttg ggaagctcgt gctctaccaa ctgagctacc accgcggaac 1440
53 ttggcggggg gtataccggc gccgctgctg agcgtcaagg acgttgccgg ttcactctca 1500
54 gtgagcacgg cgaagggtga ccagctctc gccgcggcg tctgctac cgtgtgggtg 1560

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:47

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw

```

55 ggccagtcgc gccgcgtcaa gcgtgaggac ctggacgcct acatcgcccg cgcgacggcc 1620
56 accggcgggga agcgggggtgg caaatga 1647
58 <210> SEQ ID NO: 2
59 <211> LENGTH: 548
60 <212> TYPE: PRT
61 <213> ORGANISM: Bacteriophage MX9
63 <400> SEQUENCE: 2
64 Val Ala Leu Arg Gly Ala Ser Asp Ala Thr Thr Asn Pro Ser Arg Leu
65 1 5 10 15
66 Val Gln Ser Val Ala Ala Gly Pro Arg Ala Thr Pro Trp Gly Val Ser
67 20 25 30
68 Ala Ser Trp Tyr Leu Leu Gly Arg Thr Ala Thr Gly Glu Tyr Ile Val
69 35 40 45
70 Ser Ser Asp Ala Ala Lys Lys Gly His Pro Met Ala Thr Ala Ala Glu
71 50 55 60
72 Arg Leu Pro Thr Ser Pro Ile Asp Val Asn Ala Leu Ala Leu Glu Val
73 65 70 75 80
74 Ala Arg Leu Val Ala Leu Gln Gln Gln Ser Ala Thr Pro Pro Ser Ser
75 85 90 95
76 Gly Arg Thr Phe Gly Ala Val Ala Asp Asp Trp Leu Ile Thr Glu Ala
77 100 105 110
78 Lys Arg Leu Val Cys Pro Asp Asn Glu Arg Arg His Leu Arg His Met
79 115 120 125
80 Glu Ala Leu Trp Gly Met Thr Asp Val Glu Leu Thr Pro Arg Val Val
81 130 135 140
82 Lys Ala His Leu Ala Gly Leu Leu Lys Pro Glu Gly Pro Leu Ser Ala
83 145 150 155 160
84 Ala Thr Val Asn Lys Val Arg Ser Thr Gly Lys Arg Ile Ile Lys Ala
85 165 170 175
86 Ala Gln Ile Asn Gly Glu Trp Gly Pro Val Asn Pro Phe Gly Val Leu
87 180 185 190
88 Asp Arg Glu Lys Glu Ala Lys Ala Glu Arg Leu Thr Leu Thr Ala Ala
89 195 200 205
90 Glu Cys Arg Ala Val Leu Pro His Phe Arg Ala Asp Arg Arg Arg Glu
91 210 215 220
92 Phe Leu Phe Gln Val Phe Leu Gly Pro Arg Pro Gly Glu Glu Lys Ala
93 225 230 235 240
94 Leu Leu Lys Glu Asp Val Asp Val Glu Ala Arg Thr Val Ile Phe Arg
95 245 250 255
96 Arg Ser Asn Gly Arg Asp Thr Thr Lys Thr Gly Arg Glu Arg Arg Val
97 260 265 270
98 Pro Val Pro Asp Glu Leu Trp Pro Val Leu Leu Asp Ala Met Gln Ala
99 275 280 285
100 Ser Pro Ser Asp Leu Val Phe Pro Asn Ala Lys Gly Glu Arg Gln Arg
101 290 295 300
102 Ala Asp Thr Lys Met Thr Arg Val Leu Arg Thr Ala Leu Ser Ala Ala
103 305 310 315 320
104 Gly Val Val Val Gly Trp Asp Tyr Ile Cys Arg Thr Gln Gly Cys Gly
105 325 330 335

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:47

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw

```

106 Tyr Arg Asp Val Gln Ser Gly Gly Ala Arg Gln Glu Arg Arg Cys Pro
107           340           345           350
108 Ala Cys Asp Lys Arg Met Trp Ala Ser Gly Arg Pro Lys Pro Ala Val
109           355           360           365
110 Trp Tyr Gly Leu Arg His Thr Ala Ala Thr Leu His Arg Lys Ala Gly
111           370           375           380
112 Cys Asp Pro Leu Val Ile Lys Leu Val Leu Gly His Ala Ala Val Asp
113 385           390           395           400
114 Thr Thr Asp Asp Val Tyr Thr His Leu Asp Glu Asp Tyr Cys Arg Ala
115           405           410           415
116 Glu Leu Asn Lys Leu Ser Leu Lys Ala Pro Pro Pro Pro Pro Thr His
117           420           425           430
118 Gln Gly Gly Ser Asp Gly Gly Pro Asp Ser Gly Arg Asn Thr Tyr Gly
119           435           440           445
120 Glu Gly Gly Thr Met His Gly Leu Gly Asp Leu Gln His His Arg Ala
121           450           455           460
122 Arg Ala Trp Glu Ala Arg Ala Leu Pro Thr Glu Leu Pro Pro Arg Asn
123 465           470           475           480
124 Leu Ala Gly Gly Ile Pro Ala Pro Leu Leu Ser Val Lys Asp Val Ala
125           485           490           495
126 Ala Ser Leu Ser Val Ser Thr Ala Lys Val Tyr Gln Leu Leu Ala Ala
127           500           505           510
128 Gly Val Leu Pro Thr Val Trp Val Gly Gln Ser Arg Arg Val Lys Arg
129           515           520           525
130 Glu Asp Leu Asp Ala Tyr Ile Ala Arg Ala Thr Ala Thr Gly Gly Lys
131           530           535           540
132 Arg Gly Gly Lys
133 545

```

136 &lt;210&gt; SEQ ID NO: 3

137 &lt;211&gt; LENGTH: 360

138 &lt;212&gt; TYPE: DNA

139 &lt;213&gt; ORGANISM: Bacteriophage MX9

141 &lt;400&gt; SEQUENCE: 3

```

142 gtgagctgac ctcaacggtt tgttggtggg ggagcgggac agcggaccac atggtgccag 60
143 ggcttacggc ttgcacacg gggctgggag atgctgaacg gagcgtecca tgtccacgcg 120
144 atgccgcctg gcttgacat agggattcga aacctcgacc ccgagcttgg gaagctcgtg 180
145 ctctaccaac tgagctacca ccgcaggcga agcaggggcg aaagtacggg ccgccctgtg 240
146 gcttggtcaac gggaaagtgag gtgctactcc gtctcctcga cggtagagctg gtacgagtcc 300
147 tggaagtggg actcgcggtt gcgcgcgtcc cggacctcga agaggtagac gcctggctcg 360

```

150 &lt;210&gt; SEQ ID NO: 4

151 &lt;211&gt; LENGTH: 360

152 &lt;212&gt; TYPE: DNA

153 &lt;213&gt; ORGANISM: Bacteriophage MX9

155 &lt;400&gt; SEQUENCE: 4

```

156 cgagccgggg acgggagcgg cgggaccggc ttgcgcgcgt ttacagcatc cttgctgcaa 60
157 gacgccccga ggcccgaaaa gacgaaggcc ggcagtcccg agtttctca aggactaccg 120
158 gccttcattg gtgagcggcg gaagggattc gaaccctcga ccccgagctt gggaagctcg 180
159 tgctctacca actgagctac caccgcaggc gaagcagggc gcaaagtacg ggccgccctg 240
160 tggcttggtc acgggaagtg aggtgctact ccgtctctc gacggtgagc tggtagagtc 300

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:47

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw

```

161 cctggaagtt ggactcgcggt ttgcgcgcgt cccggacctc gaagaggtag acgcctggct 360
164 <210> SEQ ID NO: 5
165 <211> LENGTH: 42
166 <212> TYPE: DNA
167 <213> ORGANISM: Bacteriophage MX9
169 <400> SEQUENCE: 5
170 gagcttggga agctcgtgct ctaccaactg agctaccacc gc 42
172 <210> SEQ ID NO: 6
173 <211> LENGTH: 240
174 <212> TYPE: DNA
175 <213> ORGANISM: Bacteriophage MX9
177 <400> SEQUENCE: 6
178 tgccagggt tacggcttcg cacacggggc tgggcgatgc tgaacggagc gtcccatgtc 60
179 cagcgcatgc cgcctggctt gcacataggg attcgaaacc tcgaccccgga gcttgggaag 120
180 ctcgccctcg acccgccag gcgttatcag ccgttcgcaa acccttactt cgccttgggg 180
181 attccggggc gggggcctgt ccatccgtcg cagcgggtag caggaggtct caggggggtt 240
184 <210> SEQ ID NO: 7
185 <211> LENGTH: 257
186 <212> TYPE: DNA
187 <213> ORGANISM: Bacteriophage MX9
189 <400> SEQUENCE: 7
190 cgccaccacc tactcaccag ggaggaagtg acggcgggcc tgactcagga cgcaacacct 60
191 acggtgaagg aggcaccatg cacggattgg gagatttgca gcatcaccgg gcgagagctt 120
192 gggaagctcg tgctctacca actgagctac caccgcggaa cttggccggg ggtataccgg 180
193 cgccgctgct gacgctcaag gacgttgccg ctctactctc agtgagcacg gcgaaggtgt 240
194 accagctcct cgccgcc 257
196 <210> SEQ ID NO: 8
197 <211> LENGTH: 20
198 <212> TYPE: DNA
199 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
202 <223> OTHER INFORMATION: Synthetic Construct
204 <400> SEQUENCE: 8
205 gaaggaggca ccatgcacgg 20
207 <210> SEQ ID NO: 9
208 <211> LENGTH: 20
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: Synthetic Construct
215 <400> SEQUENCE: 9
216 ctactgaga gtgaagccgc 20
218 <210> SEQ ID NO: 10
219 <211> LENGTH: 20
220 <212> TYPE: DNA
221 <213> ORGANISM: Artificial Sequence
223 <220> FEATURE:
224 <223> OTHER INFORMATION: Synthetic Construct
226 <400> SEQUENCE: 10

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:47

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw

```

227 cgaggtccgg gacgcgcgca 20
229 <210> SEQ ID NO: 11
230 <211> LENGTH: 19
231 <212> TYPE: DNA
232 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: Synthetic Construct
237 <400> SEQUENCE: 11
238 tgccagggct tacggcttc 19
240 <210> SEQ ID NO: 12
241 <211> LENGTH: 74
242 <212> TYPE: DNA
243 <213> ORGANISM: Myxococcus xanthus
245 <400> SEQUENCE: 12
246 gcgguuguag cucaguugu agagcacgag cuucccaagc ucggggucga gggguucgaau 60
247 ccuuccgcc gcuc 74
249 <210> SEQ ID NO: 13
250 <211> LENGTH: 20
251 <212> TYPE: DNA
252 <213> ORGANISM: Artificial Sequence
254 <220> FEATURE:
255 <223> OTHER INFORMATION: Synthetic Construct
257 <400> SEQUENCE: 13
258 tatcccagca accgccggag 20
260 <210> SEQ ID NO: 14
261 <211> LENGTH: 18
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: Synthetic Construct
268 <400> SEQUENCE: 14
269 cagcacgggt gcagcaac 18
271 <210> SEQ ID NO: 15
272 <211> LENGTH: 28
273 <212> TYPE: DNA
274 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Synthetic Construct
279 <400> SEQUENCE: 15
280 cccaattggc tcagggcagc ggctcatt 28
282 <210> SEQ ID NO: 16
283 <211> LENGTH: 31
284 <212> TYPE: DNA
285 <213> ORGANISM: Artificial Sequence
287 <220> FEATURE:
288 <223> OTHER INFORMATION: Synthetic Construct
290 <400> SEQUENCE: 16
291 ccccatggcg ctcaggggtg cgtcggacgc c 31
293 <210> SEQ ID NO: 17

```

**VERIFICATION SUMMARY**

DATE: 08/05/2004

PATENT APPLICATION: US/10/645,818

TIME: 11:12:48

Input Set : A:\Seq. Listing20547-002110.txt

Output Set: N:\CRF4\08052004\J645818.raw